

09/600766

534 Rec'd PCT/PT 21 JUL 2000

1

SEQUENCE LISTING

<110> Nabel, Gary J  
Delgado, Rafael  
Yang, Zhi-yong

<120> Targeting Gene Transfer Vectors To Certain Cell Types  
By Pseudotyping With Viral Glycoprotein

<130> 2115s-001474POA

<140> PCT/US99/01382

<141> 1999-01-21

<150> USSN 60/072033

<151> 1998-01-21

<160> 4

<170> PatentIn Ver. 2.0

<210> 1

<211> 7003

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:Plasmid  
containing DNA forGP of Ebola Virus, Ivory Coast  
strain

<220>

<221> 5'UTR

<222> (886)..(1129)

<223> CMV IE 5' UT

<220>

<221> intron

<222> (1130) .. (1840)

<223> CMV IE INT

<220>

<221> enhancer

<222> (248) .. (885)

<223> CMV Enhancer

<220>

<221> gene

<222> (1870) .. (4019)

<223> GP(IC)

<220>

<221> gene

<222> (4020) .. (4572)

<223> TbGH

<220>

<221> gene

<222> (6068) .. (6690)

<223> Kan r

<400> 1

tgcgcgcttt cggtgatgac ggtgaaaacc tctgacacat gcagctcccg gagacgggtca 60  
 cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg 120  
 ttggcgggtg tcggggctgg cttactatg cggcatcaga gcagattgta ctgagagtgc 180  
 accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240  
 ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300

tccaacatta cgcctatgtt gacattgatt attgactagt tattaatagt aatcaattac 360  
 ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaattg 420  
 cccgcctggc tgaccgcca acgacccccg ccattgacg tcaataatga cgtatgttcc 480  
 catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540  
 tgcccatctg gcagtacatc aagtgtatca tatgccaaagt acgcccccta ttgacgtcaa 600  
 tgacggtaaa tggcccgctt ggcattatgc ccagtacatg accttatggg actttcctac 660  
 ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggg tttggcagta 720  
 catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780  
 cgtcaatggg agtttgtttt ggcacaaaaa tcaacgggac tttccaaaat gtcgtaacaa 840  
 ctccgcccc a ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900  
 agtcgttcta gtgaaccgtc agatcgctg gagacgccat ccacgtgtt ttgacctcca 960  
 tagaagacac cgggaccgat ccagcctccg cggccgggaa cgggtgcattg gaacgcggat 1020  
 tccccgtgcc aagagtgcg taagtaccgc ctatagactc tataggcaca cccctttggc 1080  
 tcttatgcat gctatactgt ttttggttg gggcctatac acccccgctt ccttatgcta 1140  
 taggtgatgg tatagcttag cctataggtg tgggttattg accattattg accactcccc 1200  
 tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260  
 tattggctat atgccaatac tctgtccttc agagactgac acggactctg tatttttaca 1320  
 ggatgggggc ccatttatta ttacaaatt cacatatata acaacgcctt ccccggtgcc 1380  
 cgcagttttt attaaacata gcgtgggac tccacgcgaa tctcgggtac gtgttccgga 1440  
 catgggctct tctccggtag cggcggagct tccacatccg agccctggtc ccatgcctcc 1500  
 agcggctcat ggtcgtcgg cagctccttg ctctaacag tggaggccag acttaggcac 1560  
 agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620  
 gaaaatgagc gtggagattg ggctgcacg gctgacgcag atggaagact taaggcagcg 1680  
 gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740  
 gttgcgggtc tgtaacggg ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800  
 cgcgccacca gacataatag ctgacagact aacagactgt tcctttccat gggctctttc 1860  
 tgcagtcacc gtcgtcgaca cgtgtgatca gatatcgcg cgcgcggcc gctctagaat 1920  
 tctctaatca cagtcatcat gggagcgtca gggattctgc aattgccccg tgagcgcttc 1980  
 aggaaaacat ctttctttgt ttgggtaata atcctattcc ataaagtctt ttcaatcccc 2040  
 ttgggggttg tacacaacaa taccctacaa gtgagtgata ttgacaagtt tgtgtgccga 2100  
 gacaaactct cttaactag ccaattgaag tcagtcgggt tgaacttgga gggcaatgga 2160  
 gtagcaactg atgtaccaac ggcaacaaa agatgggggt ttcgagctgg tgttccacca 2220  
 aagggtgtaa attacgaagc tggagaatgg gctgagaact gttataacct ggctataaag 2280

aaagttagtg gtagtgagtg cctaccagaa gcccctgagg gaggtaggga ttttccccgt 2340  
 tgcgcctatg tacacaaagt ctcaggaact ggaccatgcc caggaggact cgcctttcac 2400  
 aaagaaggag cctttctcct gtatgaccga ctgcatcaa caatcattta tcgggggtaca 2460  
 agcttggcgg aaggagttat tgcatttctg atcttgccca aggcgcgaaa ggattttttc 2520  
 agtccccctt cattgcatga gcttgccaac atgaccacgg atccctccag ttactatcac 2580  
 acgacaacaa taaactacgt gggtgataat tttggaacca acaccacaga gtttctgttc 2640  
 caagtcatc atttgacgta tgtgcagctc gaggcaagat tcacaccaca attcctgttc 2700  
 ctctaaatg aaaccatcta ctctgataac cgcagaagta acacaacagg aaaactaatc 2760  
 tggaaaataa atccactgt tgataccagc atgggtgagt gggctttctg ggaaaataaa 2820  
 aaaacttcac aaaaaccctt tcaagtgaag agttgtcttt cgtacctgta ccagaaaccc 2880  
 agaaccaggt ccttgacacg acagcgacgg tctctcctcc catctccgcc cacaaccacg 2940  
 cagycgaaga ccacaaagaa ttggtttcag aggattccac tccagtgggt cagatgcaaa 3000  
 acatcaaggg aaaggacaca atgccaacca cagtgcaggg tgtaccaaca accacaccct 3060  
 ctocatttcc aatcaatgct cgcaacactg atcatacaca atcatttatc ggcttgagg 3120  
 ggccccaaaga agaccacagc accacacagc ctgccaagac caccagccaa ccaaccaaca 3180  
 gcacagaatc gacgacacta aaccaacat cagagccctc cagtagaggc acgggaccat 3240  
 ccagccccac ggtccccaac accacagaaa gccacgccga acttggaag acaaccccaa 3300  
 ccacactccc agaacagcac actgccgcca gtgccattcc aagagccgtg caccgcgacg 3360  
 aactcagtgg acctggcttc ctgacgaaca caatacgggg ggtgacaaat ctctgacag 3420  
 gatccagaag aaagcgaagg gatgtcactc ccaatacaca acccaaatgc aacccaaacc 3480  
 tgactatttg gacagccttg gatgaggggtg ctgccatagg ttagcctgg ataccatact 3540  
 tcgggccagc agctgaggga atttactctg aaggcataat ggagaatcaa aatggattga 3600  
 tctgtggatt gaggcagctg gccaacgaaa cgacacaagc tcttcaattg ttcttaaggg 3660  
 caactactga gttgcgtaca ttctctatac taaatcgga agcaatagac ttctgtctcc 3720  
 aaagatgggg aggaacatgt cacattctag ggctgattg ttgcattgaa cccaagatt 3780  
 ggacaaaaaa tatcactgat aaaattgatc aaataatcca tgactttgtc gataataatc 3840  
 ttccaaatca gaatgatggc agcaactggg ggactggatg gaaacaatgg gttcctgctg 3900  
 gaataggaat cacaggagta atcattgcta ttattgcttt gctgtgcatt tgcaaatca 3960  
 tgctttgaac taatatagca tcatacttta gaattctaga ccaggcgcct ggatccagat 4020  
 ctgctgtgcc ttctagttgc cagccatctg ttgtttgccc ctccccgtg cttccttga 4080  
 ccctggaagg tgccactccc actgtccttt cctaataaaa tgaggaaatt gcacgcatt 4140  
 gtctgagtag gtgtcattct attctggggg gtgggggtggg gcagcacagc aagggggagg 4200  
 attgggaaga caatagcagg catgctgggg atgcgggtggg ctctatgggt acccaggtgc 4260

tgaagaattg acccggttcc tcctgggcca gaaagaagca ggcacatccc cttctctgtg 4320  
 acacaccctg tccacgcccc tggttcttag ttccagcccc actcatagga cactcatagc 4380  
 tcaggagggc tccgccttca atcccacccg ctaaagtact tggagcgggc tctccctccc 4440  
 ccatcagccc accaaaccaa acctagcctc caagagtggg aagaaattaa agcaagatag 4500  
 gccattcagt gcagagggag agaaaatgcc tccaacatgt gaggaagtaa tyagagaaat 4560  
 catagaattt cttccgcttc ctgcgtcact gactcgctgc gctcggtcgt tcggctgcgg 4620  
 cgagcgggat cagctcactc aaaggcggta atacggttat ccacagaatc aggggataac 4680  
 gcaggaaaga acatgtgagc aaaaggccag caaaaggcca ggaaccgtaa aaaggccgcg 4740  
 ttgctggcgt ttttccatag gctccgcccc cctgacgagc atcacaaaaa tcgacgctca 4800  
 agtcagaggt ggcgaaaccc gacaggacta taaagatacc aggcgtttcc ccctggaagc 4860  
 tccctcgtgc gctctcctgt tccgacctg ccgcttaccg gatacctgtc cgcctttctc 4920  
 cctccgggaa gcgtggcgct ttctcaatgc tcacgctgta ggtatctcag ttcgggtgtag 4980  
 gtcgttcgct ccaagctggg ctgtgtgcac gaacccccg ttcagcccgga ccgctgcgcc 5040  
 ttatccggta actatcgtct tgagtccaac ccggttaagac acgacttatc gccactggca 5100  
 gcagccactg gtaacaggat tagcagagcg aggtatgtag gcggtgctac agagtctctg 5160  
 aagtggtagc ctaactacgg ctacactaga aggacagtat ttggtatctg cgctctgctg 5220  
 aagccagtta ccttcggaaa aagagttggg agctcttgat ccggcaaaca aaccaccgct 5280  
 ggtagcggtg gtttttttgt ttgcaagcag cagattacgc gcagaaaaaa aggatctcaa 5340  
 gaagatcctt tgatcttttc tacggggctc gacgctcagt ggaacgaaaa ctacagttaa 5400  
 gggatttttg tcatgagatt atcaaaaagg atcttcacct agatcctttt aaattaaaaa 5460  
 tgaagtttta aatcaatcta aagtatatat gagtaaactt ggtctgacag ttaccaatgc 5520  
 ttaatcagtg aggcacctat ctacgcgac tgtctatttc gttcatccat agttgcctga 5580  
 ctccgggggg gggggggcgt gaggtctgcc tcgtgaagaa ggtgttgctg actcatacca 5640  
 ggctgaatc gccccatcat ccagccagaa agtgagggag ccacggttga tgagagcttt 5700  
 gttgtaggtg gaccagttgg tgattttgaa cttttgcttt gccacggaac ggtctgcgtt 5760  
 gtcgggaaga tgcgtgatct gatccttcaa ctacgcaaaa gttcgattta ttcaacaaag 5820  
 ccgccgtccc gtcaagtcag cgtaatgctc tgccagtgtt acaaccaatt aaccaattct 5880  
 gattagaaaa actcatcgag catcaaatga aactgcaatt tattcatatc aggattatca 5940  
 ataccatatt tttgaaaaag ccgtttctgt aatgaaggag aaaactcacc gaggcagttc 6000  
 cataggatgg caagatcctg gtatcggctc gcgattccga ctctccaac atcaatacaa 6060  
 cctattaatt tcccctcgtc aaaaataagg ttatcaagt agaaatcacc atgagtgcgc 6120  
 actgaatccg gtgagaatgg caaaagctta tgcatttctt tccagacttg ttcaacaggc 6180  
 cagccattac gctcgtcatc aaaatcactc gcatcaacca aaccgttatt cattcgtgat 6240

tgcgcctgag cgagacgaaa tacgcgatcg ctgttaaaag gacaattaca aacaggaatc 6300  
 gaatgcaacc ggcgcaggaa cactgccagc gcatcaacaa tattttcacc tgaatcagga 6360  
 tatttttcta atacctggaa tgctgttttc ccgggggatcg cagtgggtgag taaccatgca 6420  
 ccatcaggag taaggataaa atgottgatg gtcggaagag gcataaattc cgtcagccag 6480  
 ttraggctga ccatctcatc tgtaacatca ttggcaacgc tacctttgcc atgtttcaga 6540  
 aacaactctg gcgcacggg cttcccatac aatcgataga ttgtcgacc tgattgccc 6600  
 acattatcgc gagcccattt ataccatata aatcagcat ccatgttgga atttaatcgc 6660  
 ggccctgagc aagacgtttc ccgttgaata tggctcataa cacccttgtt attactgttt 6720  
 atgtaagcag acagttttat tgttcatgat gatataatct tatcttgtgc aatgtaacat 6780  
 cagagatttt gagacacaac gtggctttcc ccccccccc attattgaag catttatcag 6840  
 ggttattgtc tcatgagcgg atacatattt gaatgtattt agaaaaataa acaaataagg 6900  
 gttccgcgca catttccccg aaaagtgcc cctgacgtct aagaaacat tattatcatg 6960  
 acattaacct ataaaaatag gcgtatcacg aggccctttc gtc 7003

<210> 2

<211> 7073

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:Plasmid

containing DNA for GP of Ebola Virus, Sudan strain

<220>

<221> 5'UTR

<222> (886)..(1129)

<223> CMV IE 5' UT

<220>

<221> intron

<222> (1130)..(1840)

<223> CMV IE INT

<220>

<221> enhancer

<222> (248) .. (885)

<223> CMV enhancer

<220>

<221> gene

<222> (1870) .. (4089)

<223> GP(S)

<220>

<221> gene

<222> (4090) .. (4642)

<223> TbGH

<220>

<221> gene

<222> (6138) .. (6760)

<223> Kan r

<400> 2

```
tcgcgcggttt cggatgatgac ggtgaaaacc tctgacacat gcagctcccg gagacgggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgcatggt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420
cccgcctggc tgaccgcca acgaccccg ccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttcattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600
tgacggtaaa tggccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta 720
```

catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780  
cgtcaatggg agtttgtttt ggcacaaaa tcaacgggac tttccaaaat gtcgtaacaa 840  
ctccgcccc ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900  
agctcgttta gtgaaccgtc agatcgcttg gagacgccat ccacgctgtt ttgacctcca 960  
tagaagacac cgggaccgat ccagcctccg cggccgggaa cggtgcatg gaacgcggat 1020  
tccccgtgcc aagagtgcg taagtaccgc ctatagactc tataggcaca cccctttggc 1080  
tcttatgcat gctatactgt ttttggttg gggcctatac acccccgctt ccttatgcta 1140  
taggtgatgg tatagcttag cctatagggtg tgggttattg accattattg accactcccc 1200  
tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260  
tattggctat atgccaatac tctgtccttc agagactgac acggactctg tatttttaca 1320  
ggatggggtc ccatttatta tttacaaatt cacatataca acaacgccgt cccccgtgcc 1380  
cgcagttttt attaaacata gcgtgggata tccacgcgaa tctcgggtac gtgttccgga 1440  
catgggctct tctccggtag cggcggagct tccacatccg agccctggtc ccatgcctcc 1500  
agcggctcat ggtcgctcgg cagctccttg ctctaacag tggaggccag acttaggcac 1560  
agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620  
gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680  
gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740  
gttgcggtgc tgttaacggt ggagggcagt gtagtctgag cagtactcgt tgctgccgcg 1800  
cgcgccacca gacataatag ctgacagact aacagactgt tcctttccat gggctctttc 1860  
tgcagtcacc gtcgtcgaca cgtgtgatca gatatcgcg cgcctctagc tagatgcatg 1920  
ctcgagcggc cgcagtggtg atggatatct gcagaattct atcttcagga tctcgccatg 1980  
gagggcttta gcctactcca attgcccaga gataaatttc gaaaaagctc tttctttgtt 2040  
tgggtcatca tcttatttca aaaggccttt tccatgcctt tgggtgttgt gaccaacagc 2100  
actttagaag taacagagat tgaccagcta gtctgcaagg atcatcttgc atcaactgac 2160  
cagctgaaat cagttgggtc caacctcgag gggagcggag tatctactga tatcccatct 2220  
gcgacaaagc gttggggctt cagatctgggt gtgcctcccc aagtggtcag ctatgaagca 2280  
ggagaatggg ctgaaaattg ctacaatctt gaaataaaga aaccggacgg gagcgaatgc 2340  
ttacccccac cgccggatgg tgcagagggc tttccaagggt gccgctatgt tcacaaagcc 2400  
caaggaaccg ggcctgccc gggtgactat gcctttcaca aggatggagc tttcttctc 2460  
tatgacaggg tggcttcaac tgtaatttac agaggagtca attttgctga gggggtaatc 2520  
gcattcttga tattggctaa accaaaggaa acgttccttc aatcaccccc cattcgagag 2580  
gcagcaaact acactgaaaa tacatcaagt tactatgcca catcctactt ggagtagcaa 2640  
atcgaaaatt ttggtgctca acactccacg acccttttca aaattaacaa taatactttt 2700



gttcttcttg acaggcccca cagcctcag ttccttttcc agctgaatga taccattcaa 2760  
 cttcaccaac agttgagcaa cacaactggg aaactaattt ggacactaga tgctaataatc 2820  
 aatgctgata ttggtgaatg ggcttttttg gaaaataaaa aaatctctcc gaacaactac 2880  
 gtggagaaga gctgtctttc gaaactttat cgctcaacga gacagaagac gatgatgcca 2940  
 catggtcgag aactacaaag ggaagaatct ccgaccgggc caccaggaag tattcggaac 3000  
 tggttccaaa ggattcccct gggatgggtt cattgcacgt accagaaggg gaaacaacat 3060  
 tgccgtctca gaattcgaca gaaggtcgaa gagtagatgt gaatactcag gaaactatca 3120  
 cagagacaac tgcaacaatc ataggcacta acggttaacaa catgcagatc tccaccatcg 3180  
 ggacaggact gagctccagc caaatcctga gttcctcacc gaccatggca ccaagccctg 3240  
 agactcagac ctccacaacc tacacaccaa aactaccagt gatgaccacc gaggaaccaa 3300  
 caacaccacc gagaaactct cctggctcaa caacagaagc acccactctc accaccccag 3360  
 agaataaac aacagcgggt aaaactgttt gggcacaaga gtccacaagc aacgggtctaa 3420  
 taacttcaac agtaacaggt attcttgga gccttggaact tcgaaaacgc agcagaagac 3480  
 aagttaacac cagggccacg ggtaaatgca atcccaactt aactactgg actgcacaag 3540  
 aacaacataa tgctgctggg attgcctgga tcccgactt tggaccgggt gcagaaggca 3600  
 tatacactga aggccttatg cacaaccaa atgccttagt ctgtggactc agacaacttg 3660  
 caaatgaaac aactcaagct ctgcagcttt tcttaagggc cagcagggag ctgcggacat 3720  
 ataccatact caataggaag gccatagatt tccttctgcg acgatggggc gggacatgta 3780  
 ggatcctggg accagattgt tgcattgagc cacatgattg gacaaaaac atcactgata 3840  
 aatcaacca aatcatccat gatttcacg acaaccctt acccaatcag gataatgatg 3900  
 ataattggtg gacgggctgg agacagtgga tccctgcagg aataggcatt actggaatta 3960  
 ttattgcaat cattgctctt ctttgcgtct gcaagctgct ttgttgaata tcagaattcc 4020  
 agcactggcg gccgttacta gtggatccga gctcggatcc aagctctaga ccaggcgct 4080  
 ggatocagat ctgctgtgcc ttctagttgc cagccatctg ttgtttgcc ctccccgtg 4140  
 ccttccttga ccctggaagg tgccactccc actgtccttt cctaataaaa tgaggaaatt 4200  
 gcatcgcat gtctgagtag gtgtcattct attctggggg gtgggggtgg gcagcacagc 4260  
 aagggggagg attggaaga caatagcagg catgctgggg atgcgggtgg ctctatgggt 4320  
 acccaggtgc tgaagaattg acccgttcc tcctgggcca gaaagaagca ggcacatccc 4380  
 cttctctgtg acacaccctg tccacgccc tggttcttag ttccagcccc actcatagga 4440  
 cactcatagc tcaggagggc tccgccttca atcccaccg ctaaagtact tggagcggtc 4500  
 tctccctccc tcatcagccc accaaaccaa acctagcctc caagagtggg aagaaattaa 4560  
 agcaagatag gctattaagt gcagaggag agaaaatgcc tccaacatgt gaggaagtaa 4620  
 tgagagaaat catagaattt cttccgcttc ctgctcact gactcgtgc gctcggctcg 4680

tcggctgcgg cgagcgggtat cagctcactc aaaggcggta atacggttat ccacagaatc 4740  
 aggggataac gcaggaaaga acatgtgagc aaaaggccag caaaaggcca ggaaccgtaa 4800  
 aaaggccgcg ttgctggcgt ttttccatag gctccgcccc cctgacgagc atcacaaaaa 4860  
 tcgacgctca agtcagaggt ggcgaaaccc gacaggacta taaagatacc aggcgtttcc 4920  
 ccctggaagc tccctcgtgc gctctcctgt tccgaccctg ccgcttaccg gatacctgtc 4980  
 ccgcttttctc ccttcgggaa gcgtggcgct ttctcaatgc tcacgctgta ggtatctcag 5040  
 ttcgggtgtag gtcgttcgct ccaagctggg ctgtgtgcac gaaccccccg ttcagccccg 5100  
 ccgctgcgcc ttatccggtg actatcgtct tgagtccaac ccggtgaagac acgacttattc 5160  
 gccactggca gcagccactg gtaacaggat tagcagagcg aggtatgtag gcggtgctac 5220  
 agagttcttg aagtgggtggc ctaactacgg ctacactaga aggacagtat ttggtatctg 5280  
 cgctctgctg aagccagtta ccttcggaaa aagagttggt agctcttgat ccggcaaaca 5340  
 aaccaccgct ggtagcggtg gtttttttgt ttgcaagcag cagattacgc gcagaaaaaa 5400  
 aggatctcaa gaagatcctt tgatcttttc tacgggggtct gacgctcagt ggaacgaaaa 5460  
 ctacggttaa gggattttgg tcatgagatt atcaaaaagg atcttcacct agatcctttt 5520  
 aaattaaaaa tgaagtttta aatcaatcta aagtatatat gagtaaactt ggtctgacag 5580  
 ttaccaatgc ttaatcagtg aggcacctat ctacgcgac tgtctatttc gttcatccat 5640  
 agttgcctga ctccgggggg gggggggcgt gaggtctgcc tcgtgaagaa ggtgttgctg 5700  
 actcatacca ggctgaatc gccccatcat ccagccagaa agtgagggag ccacggttga 5760  
 tgagagcttt gttgtaggtg gaccagttgg tgattttgaa cttttgcttt gccacggaac 5820  
 ggtctgcgtt gtcgggaaga tgcgtgatct gatccttcaa ctacgcaaaa gttcgattta 5880  
 ttcaacaaag ccgcggtccc gtcaagtcag cgtaatgctc tgccagtgtt acaaccaatt 5940  
 aaccaattct gattagaaaa actcatcgag catcaaatga aactgcaatt tattcatatc 6000  
 aggattatca ataccatatt ttgaaaaag ccgtttctgt aatgaaggag aaaactcacc 6060  
 gaggcagttc cataggatgg caagatcctg gtatcgggtc gcgattccga ctcgccaac 6120  
 atcaatacaa cctattaatt tcccctcgtc aaaaataagg ttatcaagt agaaatcacc 6180  
 atgagtgcag actgaatccg gtgagaatgg caaaagctta tgcatttctt tccagacttg 6240  
 ttcaacaggc cagccattac gctcgtcatc aaaatcactc gcatcaacca aaccgttatt 6300  
 cattcgtgat tgcgcctgag cgagacgaaa tacgcgatcg ctgttaaaag gacaattaca 6360  
 aacaggaatc gaatgcaacc ggcgaggaa cactgccagc gcatcaaaa tattttcacc 6420  
 tgaatcagga tattcttcta atacctggaa tgctgttttc ccggggatcg cagtgggtgag 6480  
 taaccatgca tcatcaggag tacggataaa atgcttgatg gtcggaagag gcataaatc 6540  
 cgtcagccag tttagtctga ccatctcatc tgtaacatca ttggcaacgc tacctttgcc 6600  
 atgtttcaga aacaactctg gcgcacgagg cttcccatc aatcgataga ttgtcgcacc 6660

tgattgcccg acattatcgc gagcccattt ataccatataaatcagcat ccatgttgga 6720  
 atttaatcgc ggcctcgagc aagacgtttc ccgttgaata tggctcataa cacccttgt 6780  
 attactgttt atgtaagcag acagttttat tgttcatgat gatataattt tatcttgtgc 6840  
 aatgtaacat cagagatttt gagacacaac gtggctttcc ccccccccc attattgaag 6900  
 :atttatcag ggttattgtc tcatgagcgg atacatattt gaatgtattt agaaaaataa 6960  
 acaaataggg gttccgcgca catttccccg aaaagtgcc cctgacgtct aagaaaccat 7020  
 tattatcatg acattaacct ataaaaatag gcgtatcacg aggccctttc gtc 7073

<210> 3

<211> 7285

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:Plasmid

containing DNA of GP of Ebola Virus, Zaire strain

<220>

<221> 5'UTR

<222> (886)..(1129)

<223> CMV IE 5' UT

<220>

<221> intron

<222> (1130)..(1840)

<223> CMV IE INT

<220>

<221> enhancer

<222> (248)..(885)

<223> CMV enhancer

<220>

<221> gene

<222> (1870) .. (4301)

<223> GP(Z)

<220>

<221> gene

<222> (4302) .. (4854)

<223> TbGH

<220>

<221> gene

<222> (6350) .. (6972)

<223> Kan r

<400> 3

```

tcgcgcgttt cggatgatgac ggtgaaaacc tctgacacat gcagctcccg gagacgggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300
tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt aatcaattac 360
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaatgg 420
cccgcctggc tgaccgcca acgaccccg ccattgacg tcaataatga cgtatgttcc 480
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540
tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta ttgacgtcaa 600
tgacggtaaa tggccgcct ggcattatgc ccagtacatg accttatggg actttcctac 660
ttggcagtac atctacgtat tagtcacgc tattaccatg gtgatgcggg tttggcagta 720
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc acccattga 780
cgtcaatggg agtttgtttt ggcacaaaa tcaacgggac tttccaaaat gtcgtaacaa 840
ctccgcccc ttgacgcaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900
agctcgttta gtgaaccgtc agatcgctg gagacgccat ccacgctgtt ttgacctcca 960
tagaagacac cgggaccgat ccagcctccg cggccgggaa cggtgcatg gaacgcggat 1020
tccccgtgcc aagagtgaac taagtaccgc ctatagactc tataggcaca cccctttggc 1080
tcttatgcat gctatactgt ttttggttg gggcctatac accccgctt ccttatgcta 1140

```

taggtgatgg tatagcttag cctataggtg tgggttattg accattattg accactcccc 1200  
 tattggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260  
 tattggctat atgccaatac tctgtccttc agagactgac acggactctg tattttttaca 1320  
 ggatgggggc ccatttatta ttacaaaatt cacatataca acaacgccgt ccccggtgcc 1380  
 cgcagttttc attaaacata gcgtgggacg tccacgcgaa tctcgggtac gtgttcggga 1440  
 catgggctct tctccggtag cggcggagct tccacatccg agccctggtc ccatgcctcc 1500  
 agcggctcat ggtcgtcgg cagctccttg ctccaaacag tggaggccag acttaggcac 1560  
 agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620  
 gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcagcg 1680  
 gcagaagaag atgcaggcag ctgagttggt gtattctgat aagagtcaga ggtaactccc 1740  
 gttgcggtgc tgtaaacggt ggagggcaggt gtagtctgag cagtactcgt tgctgccgcg 1800  
 cggcgcacca gacataatag ctgacagact aacagactgt tcttttccat gggctctttc 1860  
 tgcagtcacc gtctcgaca cgtgtgatca gatatcgcg cgcctctaga ccaggcgctt 1920  
 ggatcgatcc gcgatgaaga ttaagccgac agtgagcgta atcttcatct ctcttagatt 1980  
 atttgttttc cagagtaggg gtctgcaggt ccttttcaat cgtgtaacca aaataaactc 2040  
 cactagaagg atattgtggg gcaacaacac aatgggcggt acaggaatat tgcagttacc 2100  
 tctgtatcga ttcaagagga catcattctt tctttgggta attatccttt tccaaagaac 2160  
 attttccatc ccacttgag tcatccacaa tagcacatta caggtagtg atgtcgacaa 2220  
 actagtttgt cgtgacaaac tgtcatccac aaatcaattg agatcagttg gactgaatct 2280  
 cgaagggaat ggagtggcaa ctgacgtgcc atctgcaact aaaagatggg gcttcaggtc 2340  
 cgggtgtcca ccaaagggtg tcaattatga agctggtgaa tgggctgaaa actgctacaa 2400  
 tcttgaaatc aaaaaacctg acgggagtgat gtgtctacca gcagcgccag acgggattcg 2460  
 gggcttcccc cggtgccggt atgtgcacaa agtatcagga acgggaccgt gtgccggaga 2520  
 ctttgccctc cataaagagg gtgctttctt cctgtatgat cgacttgctt ccacagttat 2580  
 ctaccgagga acgactttcg ctgaagggtg cgttgcatct ctgatactgc cccaagctaa 2640  
 gaaggacttc ttcagctcac accccttgag agagccggtc aatgcaacgg aggaccgctc 2700  
 tagtggttac tattctacca caattagata tcaggctacc ggttttgga ccaatgagac 2760  
 agagtacttg ttcgaggttg acaatttgac ctacgtccaa cttgaatcaa gattcacacc 2820  
 acagtttctg ctccagctga atgagacaat atatacaagt gggaaaagga gcaataccac 2880  
 gggaaaacta atttggaagg tcaacccgga aattgatata acaatcgggg agtgggctt 2940  
 ctgggaaact aaaaaaac tcaactagaaa aattcgcagt gaagagttgt ctttcacagt 3000  
 tgtatcaaac ggagccaaaa acatcagtgg tcagagtcgg gcgcgaactt cttccgaccc 3060  
 agggaccaac acaacaactg aagaccacaa aatcatggct tcagaaaatt cctctgcaat 3120

ggttcaagtg cacagtcaag gaaggggaagc tgcaagtgtcg catctaaca cccttgccac 3180  
 aatctccacg agtccccaat ccctcacaac caaaccaggt ccggacaaca gcaccataa 3240  
 tacaccctg tataaacttg acatctctga ggcaactcaa gttgaacaac atcaccgcag 3300  
 aacagataac gacagcacag cctccgacac tccctctgcc acgaccgcag ccggaccccc 3360  
 aaaagtagag aacaccaaca cgagcaagag cactgacttc ctggaccccg ccaccacaac 3420  
 aagtcccaa aaccacagcg agaccgctgg caacaacaac actcatcacc aagataccgg 3480  
 agaagagagt gccagcagcg ggaagctagg cttaattacc aatactattg ctggagtgcg 3540  
 aggactgac acagggcgga gaagaactcg aagagaagca attgtcaatg ctcaacccaa 3600  
 atgcaaccct aatttacatt actggactac tcaggatgaa ggtgctgcaa tcggactggc 3660  
 ctggatacca tatttcgggc cagcagccga gggaatttac atagaggggc taatgcacaa 3720  
 tcaagatggt ttaatctgtg ggttgagaca gctggccaac gagacgactc aagctcttca 3780  
 actgttcctg agagccacaa ctgagctacg caccttttca atcctcaacc gtaaggcaat 3840  
 tgatttcttg ctgcagcgat ggggcggcac atgccacatt ctgggaccgg actgctgtat 3900  
 cgaaccacat gattggacca agaacataac agacaaaatt gatcagatta ttcattgattt 3960  
 tgttgataaa acccttccgg accaggggga caatgacaat tgggtggacag gatggagaca 4020  
 atggataccg gcaggtattg gagttacagg cgttataatt gcagttatcg ctttattctg 4080  
 tatatgcaa tttgtctttt agtttttctt cagattgctt catggaaaag ctcagcctca 4140  
 aatcaatgaa accaggattt aattatatgg attacttgaa tctaagatta cttgacaaat 4200  
 gataatataa tacactggag ctttaaacad agccaatgtg attctaactc ctttaaactc 4260  
 acagttaatc ataaacaagg tttggtaccg agctcgaatt atctgctgtg ccttctagtt 4320  
 gccagccatc tgttgtttgc ccctcccccg tgccttcctt gacctggaa ggtgccactc 4380  
 ccactgtcct ttcctaataa aatgaggaaa ttgcatcgca ttgtctgagt aggtgtcatt 4440  
 ctattctggg ggggtgggtg gggcagcaca gcaaggggga ggattgggaa gacaatagca 4500  
 ggcatgctgg ggatgcggtg ggctctatgg gtaccaggt gctgaagaat tgaccgggtt 4560  
 cctctgggc cagaaagaag caggcacatc ccttctctg tgacacacc tgccacgcc 4620  
 cctggttctt agttccagcc ccactcatag gacactcata gctcaggagg gctccgcctt 4680  
 caatcccacc cgctaaagta cttggagcgg totctccctc cctcatcagc ccaccaaac 4740  
 aaacctagcc tccaagagt ggaagaaatt aaagcaagat aggtatttaa gtgcagaggg 4800  
 agagaaaatg cctccaacat gtgaggaagt aatgagagaa atcatagaat ttcttccgct 4860  
 tcctcgctca ctgactcgct gcgctcggtc gttcggctgc ggcgagcgg atcagctcac 4920  
 tcaaaggcgg taatacgggt atccacagaa tcaggggata acgcaggaaa gaacatgtga 4980  
 gcaaaggcc agcaaaaggc caggaaccgt aaaaaggccg cgttgctggc gtttttccat 5040  
 aggctccgcc ccctgacga gcatcacaaa aatcgacgct caagtcagag gtggcgaaac 5100

ccgacaggac tataaagata ccaggcggtt cccctggaa gctccctcgt gcgctctcct 5160  
 gttccgaccc tgccgcttac cggatacctg tccgccttcc tcccttcggg aagcgtggcg 5220  
 ctttctcaat gctcacgctg taggtatctc agttcgggtg aggtcgttcg ctccaagctg 5280  
 ggctgtgtgc acgaaccccc cgttcagccc gaccgctgcg ccttatccgg taactatcgt 5340  
 cttgagtcca acccggttaag acacgactta tcgccactgg cagcagccac tggtaacagg 5400  
 attagcagag cgaggatatgt aggcgggtgct acagagttct tgaagtgggtg gcctaactac 5460  
 ggctacacta gaaggacagt atttgggtatc tgcgctctgc tgaagccagt taccttcgga 5520  
 aaaagagttg gtagctcttg atccggcaaa caaaccaccg ctggtagcgg tgggtttttt 5580  
 gtttgcaagc agcagattac gcgcagaaaa aaaggatctc aagaagatcc tttgatcttt 5640  
 tctacggggt ctgacgctca gtggaacgaa aactcacgtt aagggatttt ggtcatgaga 5700  
 ttatcaaaaa ggatcttcac ctagatcctt ttaaattaaa aatgaagttt taaatcaatc 5760  
 taaagtatat atgagtaaac ttggtctgac agttaccaat gcttaatcag tgaggcacct 5820  
 atctcagcga tctgtctatt tcgttcattc atagttgcct gactccgggg gggggggggc 5880  
 ctgaggtctg cctcgtgaag aaggtgttgc tgactcatac caggcctgaa tcgccccatc 5940  
 atccagccag aaagtgaggg agccacgggt gatgagagct ttgttgtagg tggaccagtt 6000  
 ggtgattttg aacttttgct ttgccacgga acggtctgcg ttgtcgggaa gatgcgtgat 6060  
 ctgatccttc aactcagcaa aagttcgatt tattcaacaa agccgcgctc ccgtcaagtc 6120  
 agcgtaatgc tctgccagtg ttacaaccaa ttaaccaatt ctgattagaa aaactcatcg 6180  
 agcatcaaat gaaactgcaa tttattcata tcaggattat caataccata tttttgaaaa 6240  
 agccgtttct gtaatgaagg agaaaactca ccgaggcagt tccataggat ggcaagatcc 6300  
 tggtatcggg ctgcgattcc gactcgtcca acatcaatac aacctattaa tttccctcgt 6360  
 tcaaaaataa ggttatcaag tgagaaatca ccatgagtga cgactgaatc cggtgagaat 6420  
 ggcaaaagct tatgcatttc tttccagact tgttcaacag gccagccatt acgctcgtca 6480  
 tcaaaatcac tcgcatcaac caaacgtta ttcatcgtg attgcgctg agcgagacga 6540  
 aatacgcgat cgctgttaaa aggacaatta caaacaggaa tcgaatgcaa ccggcgagg 6600  
 aacactgccg gcgcatcaac aatattttca cctgaatcag gatattcttc taatacctgg 6660  
 aatgctgttt tcccggggat cgcagtggtg agtaaccatg catcatcagg agtacggata 6720  
 aaatgcttga tggtcggaag aggcataaat tccgtcagcc agtttagtct gaccatctca 6780  
 tctgtaacat cattggcaac gctacctttg ccatgtttca gaaacaactc tggcgcatcg 6840  
 ggcttcccat acaatcgata gattgtcgca cctgattgcc cgacattatc gcgagcccat 6900  
 ttatacccat ataaatcagc atccatgttg gaatttaac gcggcctcga gcaagacgtt 6960  
 tcccggtgaa tatggctcat aacaccctt gtattactgt ttatgtaagc agacagtttt 7020  
 attgttcagt atgatatt tttatcttgt gcaatgtaac atcagagatt ttgagacaca 7080

acgtggcttt ccccccccc ccattattga agcatttatc agggttattg tctcatgagc 7140  
 ggatacatat ttgaatgtat ttagaaaaat aaacaaatag gggttccgcg cacatttccc 7200  
 cgaaaagtgc cacctgacgt ctaagaaacc attattatca tgacattaac ctataaaaat 7260  
 aggcgtatca cgaggccctt tcgtc 7285

<210> 4

<211> 7272

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:Plasmid  
 containing DNA for soluble GP of Ebola Virus,  
 Zaire strain

<220>

<221> 5'UTR

<222> (886)..(1129)

<223> CMV IE 5' UT

<220>

<221> intron

<222> (1130)..(1840)

<223> CMV IE INT

<220>

<221> enhancer

<222> (248)..(885)

<223> CMV enhancer

<220>

<221> gene

<222> (1870)..(4288)

<223> SGP(Z)



<220>

<221> gene

<222> (4289) .. (4841)

<223> TbGH

<220>

<221> gene

<222> (6337) .. (6959)

<223> Kan r

<400> 4

tcgcgcgttt cggatgatgac ggtgaaaacc tctgacacat gcagctcccg gagacggtca 60  
cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg 120  
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180  
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcagattgg 240  
ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata ttggctcatg 300  
tccaacatta ccgccatggt gacattgatt attgactagt tattaatagt aatcaattac 360  
ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta cggtaaattg 420  
cccgcctggc tgaccgcca acgaccccg cccattgacg tcaataatga cgtatgttcc 480  
catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt tacggtaaac 540  
tgcccacttg gcagtacatc aagtgtatca tatgccagt acgcccccta ttgacgtcaa 600  
tgacggtaaa tggcccgctt ggcattatgc ccagtacatg accttatggg actttcctac 660  
ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt tttggcagta 720  
catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc accccattga 780  
cgtcaatggg agtttggttt ggcacaaaaa tcaacgggac tttccaaaat gtcgtaacaa 840  
ctccgccccca ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct atataagcag 900  
agctcgttta gtgaaccgct agatcgcttg gagacgcat ccacgtggt ttgacctcca 960  
tagaagacac cgggaccgat ccagcctccg cggccgggaa cggatgcattg gaacgcggat 1020  
tccccgtgcc aagagtgacg taagtaccgc ctatagactc tataggcaca cccctttggc 1080  
tcttatgcat gctatactgt ttttggttg gggcctatac acccccgctt ccttatgcta 1140  
taggtgatgg tatagcttag cctatagggt tgggttattg accattattg accactcccc 1200  
tattgggtgac gatactttcc attactaatc cataacatgg ctctttgcca caactatctc 1260  
tattggctat atgccaatac tctgtccttc agagactgac acggactctg tatttttaca 1320  
ggtgggggtc ccatttatta tttacaaatt cacatatata acaacgccgt ccccgtgcc 1380

cgcagttttt attaaacata gcgtgggata tccacgcgaa tctcgggtac gtgttccgga 1440  
 catgggctct tctccggtag cggcggagct tccacatccg agccctggtc ccatgcctcc 1500  
 agcggctcat ggtcgctcgg cagctccttg ctcctaacag tggaggccag acttaggcac 1560  
 agcacaatgc ccaccaccac cagtgtgccg cacaaggccg tggcggtagg gtatgtgtct 1620  
 gaaaatgagc gtggagattg ggctcgcacg gctgacgcag atggaagact taaggcajcg 1680  
 gcagaagaag atgcaggcag ctgagttgtt gtattctgat aagagtcaga ggtaactccc 1740  
 gttgcgggtgc tgttaacggg ggagggcagt gtagtctgag cagtactcgt tgctgcccg 1800  
 cgcgccacca gacataatag ctgacagact aacagactgt tcctttccat gggctttttc 1860  
 tgcagtcacc gtcgtcgaca cgtgtgatca gatatcgagg ccgctctaga ccaggcgcct 1920  
 ggatcgaatt gatgaagatt aagccgacag tgagcgtaat cttcatctct cttagattat 1980  
 ttgttttcca gagtaggggt cgtcagggtcc ttttcaatcg tgtaaccaa ataaactcca 2040  
 ctagaaggat attgtggggc aacaacacaa tgggcgttac aggaatattg cagttacctc 2100  
 gtgatcgatt caagaggaca tcattctttc tttgggtaat tatccttttc caaagaacat 2160  
 tttccatccc acttgagtc atccacaata gcacattaca ggtagtgat gtcgacaaac 2220  
 tagtttgctg tgacaaactg tcatccacaa atcaattgag atcagttgga ctgaatctcg 2280  
 aagggaatgg agtggaact gacgtgccat ctgcaactaa aagatggggc ttcagggtccg 2340  
 gtgtcccacc aaagggtggtc aattatgaag ctggtgaatg ggctgaaaac tgctacaatc 2400  
 ttgaaatcaa aaaacctgac gggagtgagt gtctaccagc agcgcagac gggattcggg 2460  
 gcttcccccg gtgccggtat gtgcacaaag tatcaggaac gggaccgtgt gccggagact 2520  
 ttgccttcca taaagagggt gctttcttcc tgtatgatcg acttgcttcc acagttatct 2580  
 accgaggaac gactttcgct gaagggtgctg ttgcatttct gatactgccc caagctaaga 2640  
 aggacttctt cagctcacac cccttgagag agccgggtcaa tgcaacggag gaccgtctta 2700  
 gtggctacta ttctaccaca attagatata aggtaccgg ttttgaacc aatgagacag 2760  
 agtacttggt cgagggtgac aatttgacct acgtccaact tgaatcaaga ttcacaccac 2820  
 agtttctgct ccagctgaat gagacaatat atacaagtgg gaaaaggagc aataccacgg 2880  
 gaaaactaat ttggaaggtc aaccccgaaa ttgatacaac aatcggggag tgggccttct 2940  
 gggaaactaa aaaaacctca ctagaaaaat tcgcagtga gagttgtctt tcacagttgt 3000  
 atcaaacgga gccaaaaaca tcagtgggtca gagtccggcg cgaacttctt ccgaccagg 3060  
 gaccaacaca acaactgaag accacaaaat catgggttca gaaaattcct ctgcaatggt 3120  
 tcaagtgcac agtcaaggaa gggaagctgc agtgtcgcat ctaacaaccc ttgccacaat 3180  
 ctccacgagt cccaatccc tcacaaccaa accagggtccg gacaacagca ccataatac 3240  
 acccgtgtat aaacttgaca tctctgaggc aactcaagtt gaacaacatc accgcagaac 3300  
 agacaacgac agcacagcct ccgacactcc ctctgccacg accgcagccg gaccccaaa 3360

agcagagaac accaacacga gcaagagcac tgacttctg gaccccgcca ccacaacaag 3420  
 tccccaaaac cacagcgaga ccgctggcaa caacaacact catcaccaag ataccggaga 3480  
 agagagtgcc agcagcgga agctaggctt aattaccaat actattgctg gagtcgcagg 3540  
 actgatcaca ggcgggagaa gaactcgaag agaagcaatt gtcaatgctc aacccaaatg 3600  
 caaccctaatt ttacattact ggactactca ggatgaagggt gctgcaatcg gactggcctg 3660  
 gataccatat ttccggccag cagccgaggg aatttacata gaggggctaa tgcacaatca 3720  
 agatgggttta atctgtgggt tgagacagct ggccaacgag acgactcaag ctcttcaact 3780  
 gttctcgaga gccacaactg agctacgcac cttttcaatc ctcaaccgta aggcaattga 3840  
 tttcttgctg cagcgatggg gcggcacatg ccacattctg ggaccggact gctgtatcga 3900  
 accacatgat tggaccaaga acataacaga caaaattgat cagattattc atgattttgt 3960  
 tgataaaacc cttccggacc aggggggaaa tgacaattgg tggacaggat ggagacaatg 4020  
 gataccggca ggtattggag ttacaggcgt tataattgca gttatcgctt tattctgtat 4080  
 atgcaaattt gtcttttagt ttttcttcag attgcttcat ggaaaagctc agcctcaaatt 4140  
 caatgaaacc aggatttaatt tatatggatt acttgaatct aagattactt gacaaatgat 4200  
 aatataatac actggagctt taaacatagc caatgtgatt ctaactcctt taaactcaca 4260  
 gttaatcata aacaagggtt ggaattgatc tgctgtgcct tctagttgcc agccatctgt 4320  
 tgtttgcccc tccccgtgc ctctcttgac cctggaagggt gccactccca ctgtcctttc 4380  
 ctaataaaat gaggaattg catcgcattg tctgagtagg tgtcattcta ttctgggggg 4440  
 tggggtgggg cagcacagca agggggagga ttgggaagac aatagcaggc atgctgggga 4500  
 tgcggtgggc tctatgggta cccagggtgt gaagaattga cccgggttct cctgggccag 4560  
 aaagaagcag gcacatcccc ttctctgtga cacaccctgt ccacgcccct ggttcttagt 4620  
 tccagcccca ctcataggac actcatagct caggaggggt ccgccttcaa tcccaccgcg 4680  
 taaagtactt ggagcggctt ctccctcct catcagccca ccaaaccaaa cctagcctcc 4740  
 aagagtggga agaaattaaa gcaagatagg ctattaagt cagagggaga gaaaatgcct 4800  
 ccaacatgtg aggaagtaat gagagaaatc atagaatttc ttccgcttcc tcgctcactg 4860  
 actcgtgcg ctcggtcgtt cggctgcggc gagcggatc agtcactca aaggcggtaa 4920  
 tacggttatc cacagaatca ggggataacg caggaaagaa catgtgagca aaaggccagc 4980  
 aaaaggccag gaaccgtaaa aaggccgctg tgctggcgtt tttccatagg ctccgcccc 5040  
 ctgacgagca tcacaaaaat cgacgtcaa gtcagagggt gcgaaaccg acaggactat 5100  
 aaagatacca ggcgtttccc cctggaagct cctcgtgcg ctctcctgtt ccgaccctgc 5160  
 cgcttaccgg atacctgtcc gcctttctcc cttcgggaag cgtggcgctt tctcaatgct 5220  
 cacgtgtag gtatctcagt tcggtgtagg tcgttcgctc caagctgggc tgtgtgcacg 5280  
 aacccccgt tcagcccgac cgctgcgctt tatccggtaa ctatcgtctt gagtccaacc 5340

cggttaagaca cgacttatcg ccaactggcag cagccactgg taacaggatt agcagagcga 5400  
 ggtatgtagg cggtgctaca gagttcttga agtgggtggc taactacggc tacactagaa 5460  
 ggacagtatt tggatatctgc gctctgctga agccagttac cttcggaaaa agagttggta 5520  
 gctcttgatc cggaacaaa accaccgctg gtagcgggtg tttttttgtt tgcaagcagc 5580  
 agattacggc cagaaaaaaa ggatctcaag aagatccttt gatcttttct acgggggtctg 5640  
 acgctcagtg gaacgaaaac tcacgttaag ggatttttgt catgagatta tcaaaaagga 5700  
 tcttcaccta gatcctttta aattaaaaat gaagttttta atcaatctaa agtatatatg 5760  
 agtaaacttg gtctgacagt taccaatgct taatcagtg ggcacctatc tcagcgatct 5820  
 gtctatctcg ttcattccata gttgctgac tccggggggg gggggcgctg aggtctgcct 5880  
 cgtgaagaag gtgttgctga ctcataccag gcctgaatcg ccccatcatc cagccagaaa 5940  
 gtgagggagc caggttgat gagagctttg ttgtaggtgg accagtttgt gattttgaac 6000  
 ttttgctttg ccacggaacg gtctgcgttg tccgggaagat gcgtgatctg atccttcaac 6060  
 tcagcaaaag ttcgatttat tcaacaaagc cgccgtcccg tcaagtcagc gtaatgctct 6120  
 gccagtgtta caaccaatta accaattctg attagaaaaa ctcatcgagc atcaaatgaa 6180  
 actgcaattt attcatatca ggattatcaa taccatattt ttgaaaaagc cgtttctgta 6240  
 atgaaggaga aaactcaccg aggcagttcc ataggatggc aagatcctgg tatcgggtctg 6300  
 cgattccgac tcgtccaaca tcaatacaac ctattaattt cccctcgta aaaataaggt 6360  
 tatcaagtga gaaatcacca tgagtgcga ctgaatccgg tgagaatggc aaaagcttat 6420  
 gcatttcttt ccagacttgt tcaacaggcc agccattacg ctcgatcatc aaatcactcg 6480  
 catcaaccaa accgttattc attcgtgatt gcgcctgagc gagacgaaat acgcgatcgc 6540  
 tgttaaaagg acaattacaa acaggaatcg aatgcaaccg gcgcaggaac actgccagcg 6600  
 catcaacaat attttcacct gaatcaggat attcttctaa tacctggaat gctgttttcc 6660  
 cggggatcgc agtgggtgagt aaccatgcat catcaggagt acggataaaa tgcttgatgg 6720  
 tcggaagagg cataaattcc gtcagccagt ttagtctgac catctcatct gtaacatcat 6780  
 tggcaacgct acctttgcca tgtttcagaa acaactctgg cgcacgggc tccccatata 6840  
 atcgatagat tgtcgcacct gattgccga cattatcgcg agccattta taccatata 6900  
 aatcagcatc catgttgga tttaatcgcg gcctcgagca agacgtttcc cgttgaatat 6960  
 gggtcataac accccttgta ttactgttta tgtaagcaga cagttttatt gttcatgatg 7020  
 atatattttt atcttgatga atgtaacatc agagattttg agacacaacg tggctttccc 7080  
 cccccccca ttattgaagc atttatcagg gttattgtct catgagcgga tacatatttg 7140  
 aatgtattta gaaaaataaa caaatagggg ttccgcgcac atttccccga aaagtgccac 7200  
 ctgacgtcta agaaaccatt attatcatga cattaaccta taaaaatagg cgtatcacga 7260  
 ggccctttcg tc